

# AI PROJECT CYCLE



# Imagine:



- The world's largest diamond, is in danger as Mr. X has threatened to steal it.
- No one is able to track Mr. X and so the situation is critical.
- You have been appointed as the Chief Security Officer and your job is to enhance the security of the diamond to make the area impossible for Mr. X to break into and steal the diamond.

# The Aim of this System is to:

- Secure the Diamond
  - As you interacted with the authorities, you get to know that some people are allowed to enter the area where the diamond is kept. Some of them being - the maintenance people; officials; VIPs, etc.
  - Now, the challenge is to make sure that no unauthorized person enters the premises.

# Problem Scoping



- While finalizing the aim of this system, you scope the problem which you wish to solve with the help of your project. This is **Problem Scoping**.



- For the Purpose you need to decide on one:
  - a. Get photographs of all the authorised people.
  - b. Get photographs of all the unauthorised people.
  - c. Get photographs of the premises in which the diamond has been kept.
  - d. Get photographs of all the visitors.

# Data Acquisition:

- As you start collecting the photographs, you actually acquire data in a visual form. This data now becomes the base of your security system. Note that the data needs to be accurate and reliable as it ensures the efficiency of your system. This is known as **Data Acquisition**.





- **You Realise:**

- *After acquiring the required data -*

- That it is not uniform.
    - Some images are small in size while the others are big.
    - Some images are missing while you have multiple copies of others.
    - Thus, you think of putting all the information collected in a simplified format.



- **For Which:**
  - a. Create a table and put the names of people whose photographs you have.
  - b. Put all the photographs in a graph and try to interpret a pattern out of it.
  - c. Make a database to store the image data.
  - d. Remember all the faces you see in the images.



# Data Exploration



- At this stage, you try to interpret some useful information out of the data you have acquired.
- For this, you explore the data and try to put it uniformly for a better understanding. This is known as **Data Exploration**.



- Now you need to develop a System:
  - Which detects the face of a person entering the vault and match it with the existing image data you have in your system.
  - For this, you put all your data into the AI-enabled model and train it in such a way that it alerts the officials if an unknown person tries to enter the vault.

# Modelling



- To implement your idea, you now look at different AI-enabled algorithms which work on Computer Vision (since you are working on visual data).
- You go through several models and select the ones which match your requirements.
- After choosing the model, you implement it. This is known as the Modelling stage.

# Test Time



- Your surveillance system is now complete!
- You test it by sending a mix of known and unknown faces to the vault. You notice that the results were 70% correct.
- After evaluating this model, you work on other shortlisted AI algorithms and work on them.

# Evaluation



- As you move towards deploying your model in the real-world, you test it in as many ways as possible. The stage of testing the models is known as **Evaluation**.
- In this stage, we evaluate each and every model tried and choose the model which gives the most efficient and reliable results.
- After proper testing, we deploy our surveillance system in the premises. Mr. X, who is unaware of the surveillance system, tries to break through the vault and gets caught in your system. We have saved the diamond!

# AI Project Cycle - Defined!



- Start with **Problem Scoping**, set the goal for the AI project by stating the problem which you wish to solve with it.
- Acquire data which will become the base of the project as it will help us in understanding what are the parameters that are related to the problem we are scoping.
- Go for **data acquisition** by collecting data from various reliable and authentic sources.
- Since the data we collect would be in large quantities, you can try to give it a visual form of different types of representations like graphs, databases, flow charts, maps, etc. This makes it easier for us to



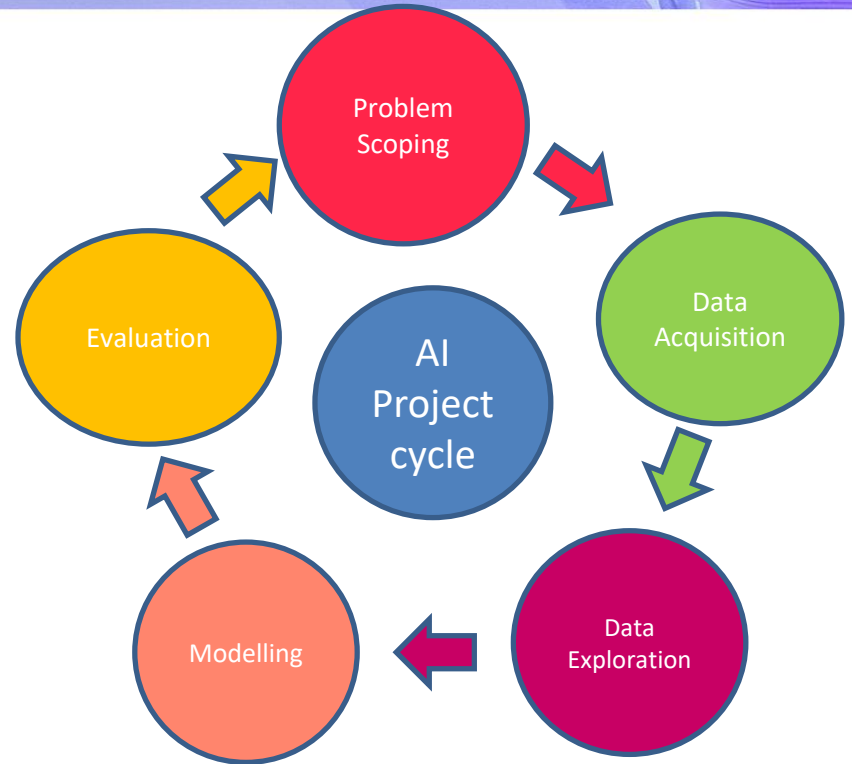
# AI Project Cycle - Defined!



- After exploring the patterns, we can decide upon the type of model we would build to achieve the goal. For this, we can research online and **select various models** which give a suitable output.
- We can test the selected models and figure out which is the most efficient one.
- The most efficient model is now the base of our AI project and we can develop our algorithm around it.
- Once the modelling is complete, we now need to **test our model** on some newly fetched data. The results will help us in evaluating our model and hence improving it.
- **Finally, after evaluation**, the project cycle is now complete and what we get is our AI project.

# AI PROJECT CYCLE

- Life cycle of AI project
- **Project Cycle** is a step-by-step **process to solve** problems using **proven scientific methods** and drawing inferences about them.

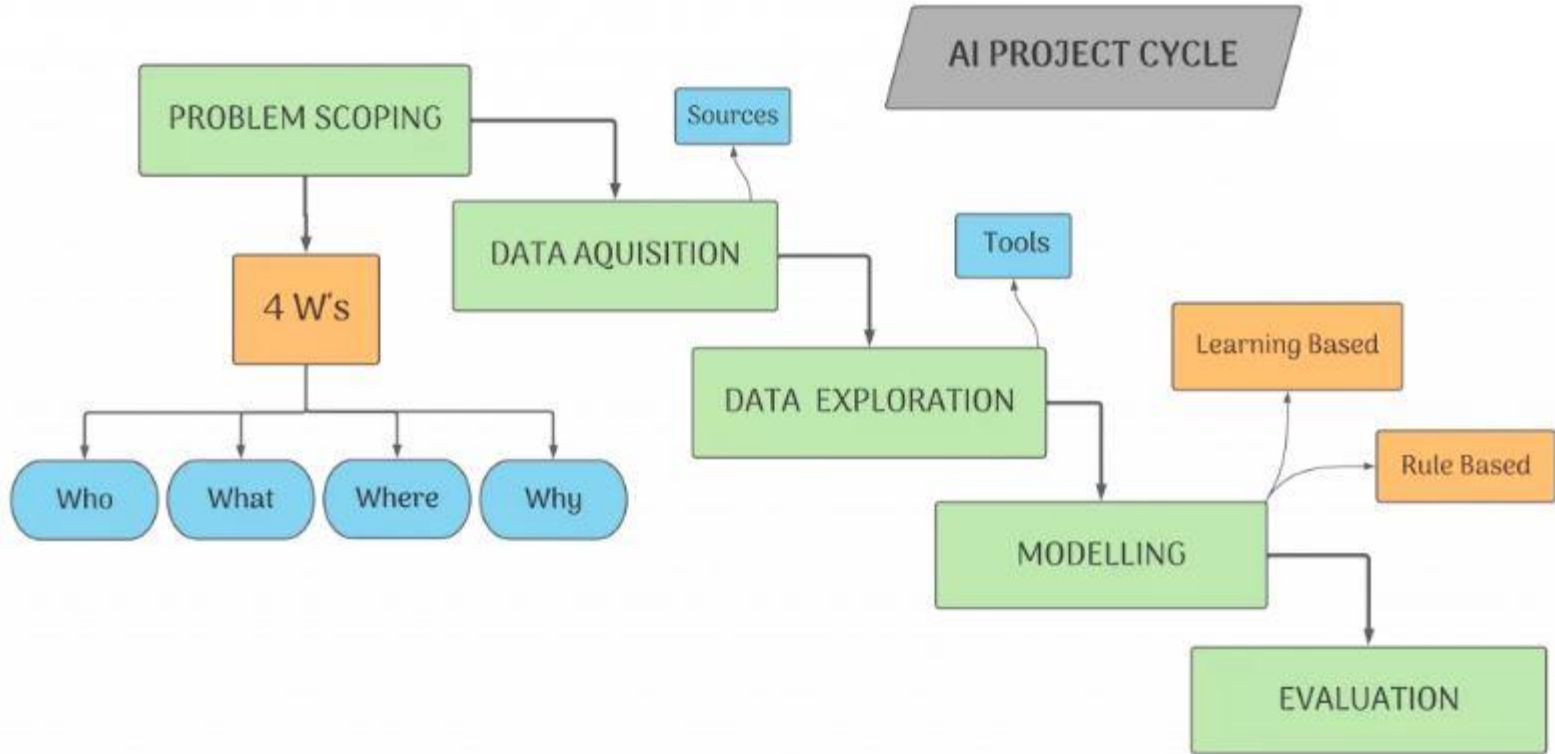


# AI PROJECT CYCLE



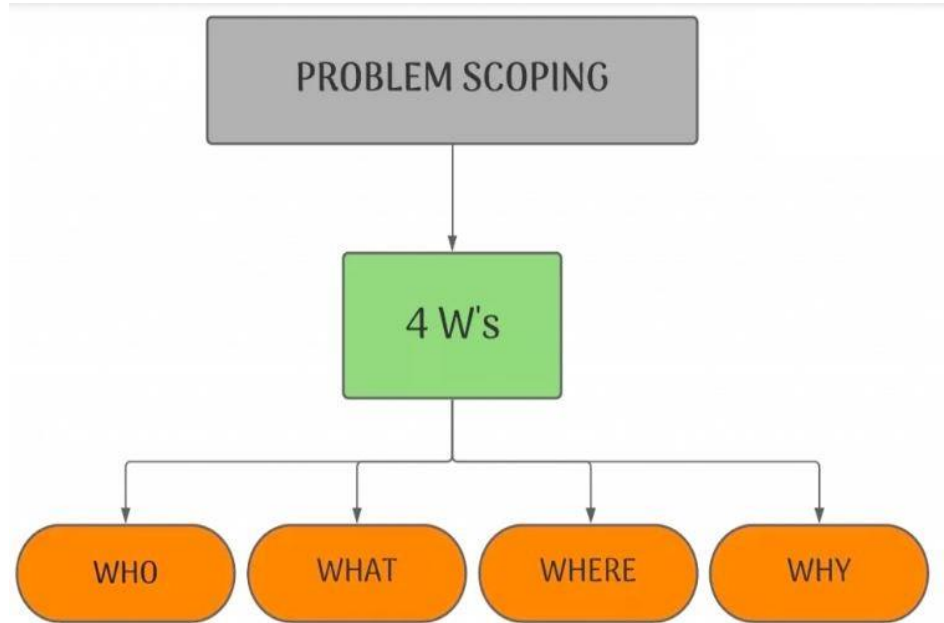
<b>Problem Scoping</b>	Understanding the problem
<b>Data Acquisition</b>	Collecting accurate and reliable data
<b>Data Exploration</b>	Arranging the data uniformly
<b>Modelling</b>	Creating Models from the data
<b>Evaluation</b>	Evaluating the project

# AI PROJECT CYCLE



# Problem Scoping:

- It is a fact that we are surrounded by problems.
- They could be small or big, sometimes ignored or sometimes even critical.
- Many times we become so used to a problem that it becomes a part of our life.
- Identifying such a problem and having a vision to solve it, is what Problem Scoping is about.





# Problem Scoping

- **Who?** - Refers that who is facing a problem and who are the stakeholders of the problem.
- **What?** - Refers to what is the problem and how you know about the problem.
- **Where?** - It is related to the context or situation or location of the problem.
- **Why?** - Refers to why we need to solve the problem and what are the benefits to the stakeholders after solving the problem.

The final outcome of problem scoping is the problem statement template.





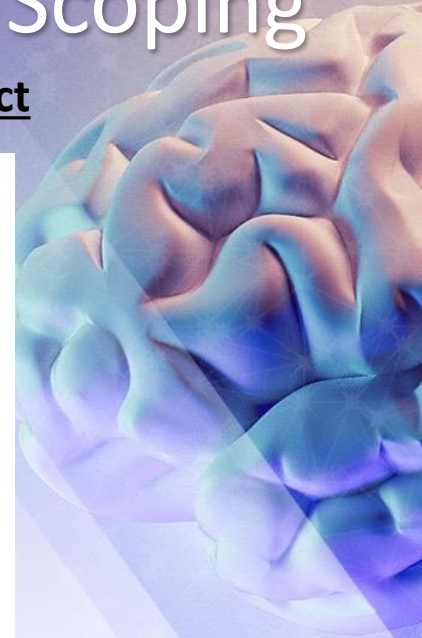
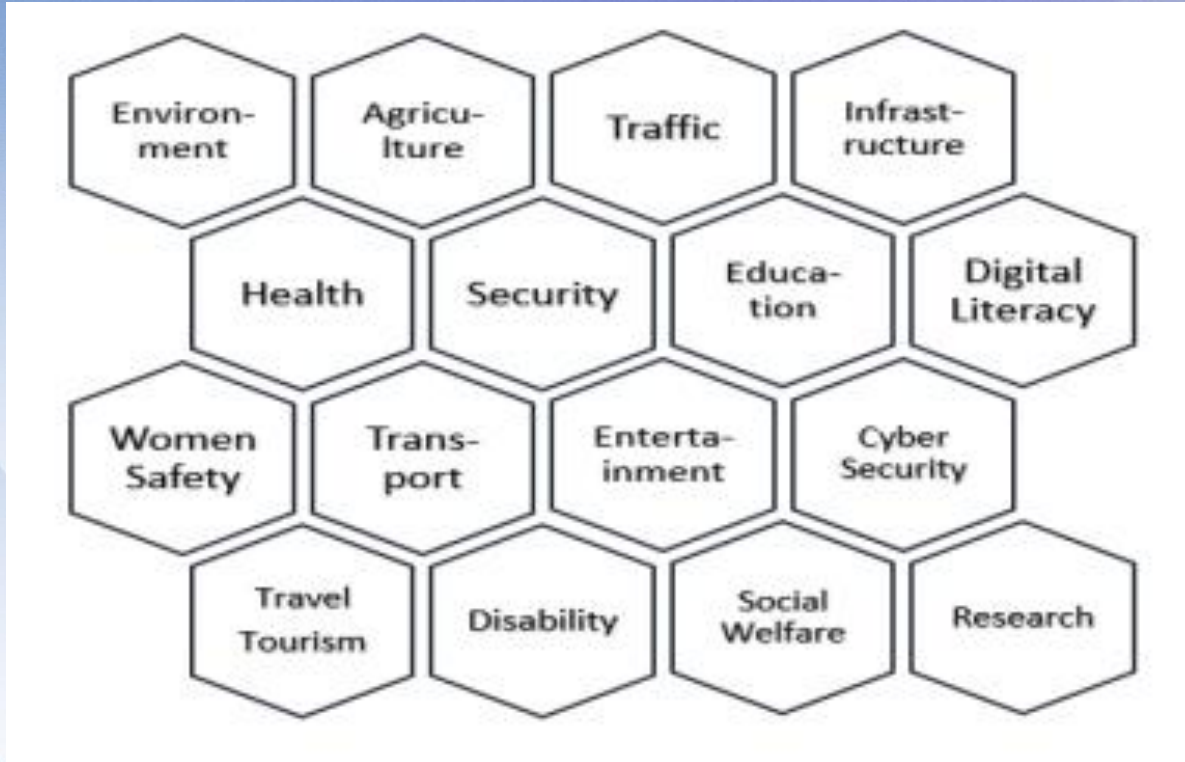
# The problem statement template

- When the above 4Ws are completely filled you need to prepare a summary of these 4Ws.
- This summary is known as the problem statement template.
- This template explains all the key points in a single template.
- So if the same problem arises in the future this statement helps to resolve it easily.



# Problem Scoping

Activity - Brainstorm around the theme and set a goal for the AI project



# Problem Scoping:

- Select a theme

Agriculture

1. Yield rates
2. Sowing and harvesting patterns

Digital literacy

1. Online learning platforms
2. Digital awareness
3. Ebooks

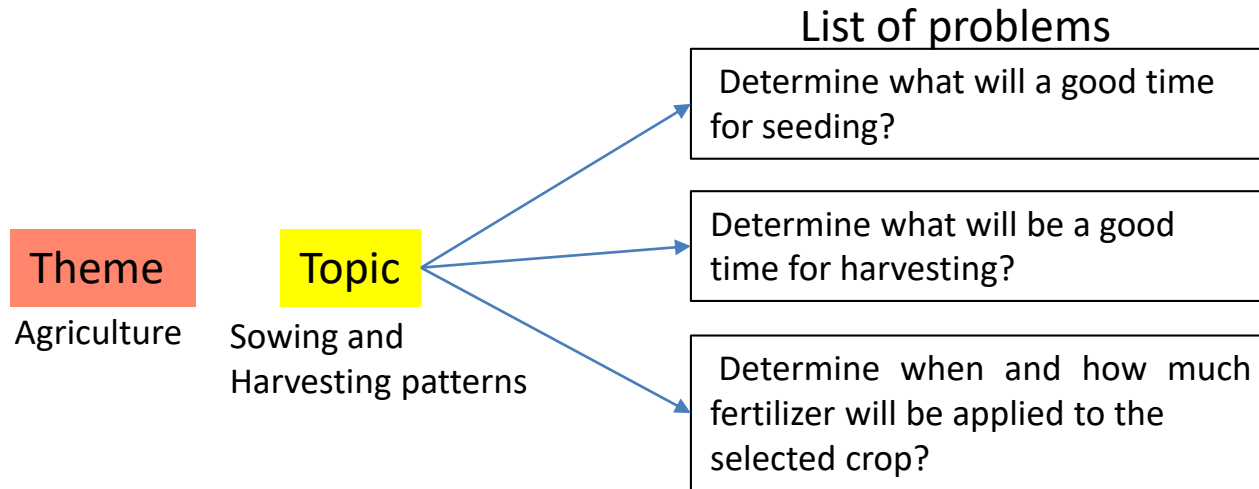
Health

1. Medicinal Aid
2. Mobile medications
3. Spreading of diseases

# Problem Scoping

## 1. Select a theme

Eg: Agriculture: this theme includes different topics like pest issues, sowing and harvesting patterns, yield rates etc.



# Problem Scoping:4Ws Problem Canvas:



## ***Who?***

- The “Who” block helps you in analyzing the people getting affected directly or indirectly due to it.
- Under this, you find out who the ‘Stakeholders’ to this problem are and what you know about them.
- Stakeholders are the people who face this problem and would be benefited with the solution.



# Problem Scoping:4Ws Problem Canvas:



## Who

- *Que.* Who are the Stakeholders?
- *Ans.* Farmers, Fertilizer producers, Labours and Tractor companies.
  
- *Que.* What do you know about them?
- *Ans.* These are some of the peoples worst affected by the problem and loses their money and time.



# Problem Scoping:4Ws Problem Canvas:



## *What?*

- Under the “What” block, you need to look into what you have on hand.
- At this stage, you need to determine the nature of the problem.
- What is the problem and how do you know that it is a problem?
- Under this block, you also gather evidence to prove that the problem you have selected actually exists.
- Newspaper articles, Media, announcements, etc are some examples.

# Problem Scoping:4Ws Problem Canvas:

- **What**
- *Que.* What is the problem?
- *Ans.* Determine what will be the best time for seeding or crop harvesting.
- *Que.* How do you know that it is a problem? (Is there any evidence?)
- *Ans.* Seeding at improper position and imperfect time leads to the wastage of money and time.

# Problem Scoping:4Ws Problem Canvas:



- ***Where?***
- Now that you know who is associated with the problem and what the problem actually is; you need to focus on the context/situation/location of the problem.
- This block will help you look into the situation in which the problem arises, the context of it, and the locations where it is prominent.

# Problem Scoping:4Ws Problem Canvas:

- **Where**

- *Que.* What is the context/situation in which the stakeholders experience the problem?
- *Ans.* Decide the mature age for the crop and determine its time.
- *Que.* Where is the problem located?
- *Ans.* In the field.

# Problem Scoping:4Ws Problem Canvas:

- **Why?**
- You have finally listed down all the major elements that affect the problem directly.
- Now it is convenient to understand who the people that would be benefitted by the solution are; what is to be solved; and where will the solution be deployed.
- These three canvases now become the base of why you want to solve this problem.
- Thus in the “Why” canvas, think about the benefits which the stakeholders would get from the solution and how would it benefit them as well as the society.

# Problem Scoping:4Ws Problem Canvas:

- **Why**
- *Que.* Why will this solution be of value to the stakeholders?
- *Ans.* This solution will help us determine the exact time for sowing and harvesting of crops, thus leading to the maximum yield.
  
- *Que.* How will the solution improve their situation?
- *Ans.* They could maximize their profit.

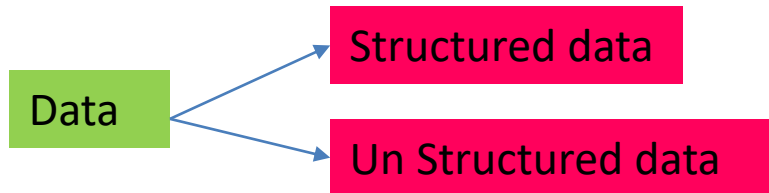


# Final Problem statement

	<b>Stakeholders</b>	<b>Who</b>
Our	Farmers, Fertilizer Producers, Labours, Tractor Companies	
has /have a problem that	The problem, Issue, Need	What
	Determine what will a good time for seeding or crop harvesting?	
When/while	Context/Situation	When
	Decide the mature age for the crop and determine its time	
An ideal solution would	Solution Benefits	Why
	Take the crop on time and supply against market demand on time	

# Data Acquisition

- **Data:**
- Data can be defined as a representation of facts or instructions about some entity (students, school, sports, business, animals etc.) that can be processed or communicated by human or machines.
- Data is a collection of facts, such as numbers, words, pictures, audio clips, videos, maps, measurements, observations or even just descriptions of things.
- Data may be represented with the help of characters such as alphabets (A-Z, a-z), digits (0-9) or special characters (+, -, /, \*, <, >, = etc.)



# Data Acquisition



## Structured Data:

- Structured data is categorized as quantitative data.
- It's the type of data most of us work with every day.
- Structured data has predefined data types and format so that it fits well in the column/fields of database or spreadsheet.
- They are highly organised and easily analysed.
- The data is structured in accurately defined fields.
- The data is stored in relational databases or spreadsheets (like Excel).
- Examples of Structured data are: Name, age, address, cricket score board, school time table etc.

# Data Acquisition



## Unstructured Data:

- Unstructured data is categorized as qualitative data.
- It cannot be processed and analysed using conventional relational database (RDBMS) methods.
- Unstructured data is difficult to deconstruct because it has no predefined model, meaning it cannot be organized in relational databases.
- Instead, non relational, or NOSQL databases, are best fit for managing unstructured data.
- Examples of unstructured data include text, video, audio, mobile activity, social media activity, satellite imagery, surveillance imagery and the list goes on.

# Data Acquisition



- **Datasets:**
- A Data set is a set or collection of data.
- This set is normally presented in a tabular pattern.
- Every column describes a particular variable and each row corresponds to a given member of the data set.
- Data sets describe values for each variable for unknown quantities such as height, weight, temperature, volume, etc of an object or values of random numbers. The values in this set are known as a datum.
- The data set consists of data of one or more members corresponding to each row.

# Data Acquisition



- Data set eg:

Name of Students	Attendance (in %)	Gender	Participation in Sports	Total Marks obtained (in %)
A	76	FEMALE	Y	92
B	82	MALE	Y	88
C	84	MALE	N	87
D	92	FEMALE	N	90
E	78	MALE	Y	86
F	67	FEMALE	Y	80
G	90	FEMALE	N	88



# Data Acquisition



Training data	Test Data
<p>A training dataset is a database of examples used during the learning process and is used to fit the parameters.</p>	<p>A test set is set of example used only to access the performance the fully specified classifier.</p>
<p>Maximum part of the dataset comes under training data (Usually 80%).</p>	<p>A very little part of dataset is used for test data (Usually 20%).</p>

**Note:** The training data and the test data are not different, they are usually divided from the main dataset in 80% - 20%.

# Data Acquisition



## Data Features:

- Data features refer to the type of data we want to collect.
- A measurable piece of data that can be used for analysis.
- Features are also sometimes referred to as “variables” or “attributes.”
- Depending on what we're trying to analyze, the features we include in our dataset can vary widely.
- In CSV, Excel and databases, they could be seen as columns.

# Data Acquisition

- Data sources

Surveys

Web Scraping

Sensors

Cameras

Observations

API(Application  
Program  
Interface)

# Data Acquisition



## Acquiring Data:

- **Surveys:**
  - A research method used for collecting data from a predefined group of respondents to gain information and insights into various topics of interest.
- **Camera:**
  - A camera captures a visual image.
    - A device for recording visual images in the form of photographs, film, or video signals.
  - Could be used to collect data for CV projects.
- **Web Scraping:**
  - Web scraping is the process of collecting structured web data in an automated fashion. It's also called web data extraction.
  - Some of the main use cases of web scraping include price monitoring, price intelligence, news monitoring, lead generation, and market research among many others.

# Data Acquisition



- Observations:
  - Some data we can acquire through monitoring and close inspection.
- **Sensors:**
  - A device which detects or measures a physical property and records, indicates, or otherwise responds to it.
  - *E.g.* - Temperature Sensors, Humidity Sensors, Pressure Sensors, Proximity Sensors, Level Sensors, Accelerometers, Gyroscope, Infrared Sensors, etc.
- **API (Application Program Interface):**
  - An API is a software intermediary that allows two applications to talk to each other.

# Data Acquisition



## Acquiring Data from Reliable Sources:

- Sometimes, we use the internet and try to acquire data for your project from some random websites.
- Such data might not be authentic as its accuracy cannot be proved.
- Due to this, it becomes necessary to find a reliable source of data from where some authentic information can be taken.
- At the same time, we should keep in mind that the data which we collect is open-sourced and not someone's property.



# Data Acquisition



## System maps

- System maps help us to find the relationship between the elements of the problem which we have scoped
- A system map shows the components and boundary of a system and the components of the environment at a point in time.
- The main use of a system map is to help structure a system and communicate the result to others.
- It helps us in strategizing the solution for achieving the goal of our project.
- help to understand complex issues with multiple factors that affect each other
- **Circles** represents the elements,
- **Arrows** represents the relationship between the elements. Length of arrow represents time for a change to happen. This is **time delay**. **The arrow-head** depicts the direction of the effect and the sign (+ or -) shows their relationship. **If the arrow goes from X to Y with a + sign, it means that both are directly related to each other. If the arrow goes from X to Y with a – sign, it means that both the elements are inversely related to each other .**
- **Loops** represent a specific chain of causes and effects.
- To change the outcome of a system, as a change maker, either change the elements in a system or change the relationships between elements.

# Data Exploration:



- Arrange the gathered data uniformly for a better understanding.
- A process of interpreting the acquired data to find out useful information from the acquired data.
- The goal of data exploration is to learn about characteristics and potential problems of a data set without the need to formulate assumptions about the data beforehand.
- In statistics, data exploration is often referred to as “exploratory data analysis” and contrasts traditional hypothesis testing.
- It was most prominently promoted by the statistician “John W. Tukey” to suggest hypotheses about the causes of observed phenomena and to assess assumptions on which statistical methods are based.
- Since its beginnings, EDA has been a very graphical approach.
- Typical plots include histograms, box plots, scatter plots and many more in order to learn about distributions, correlations, outliers, trends, and other data characteristics.

# Data Exploration:



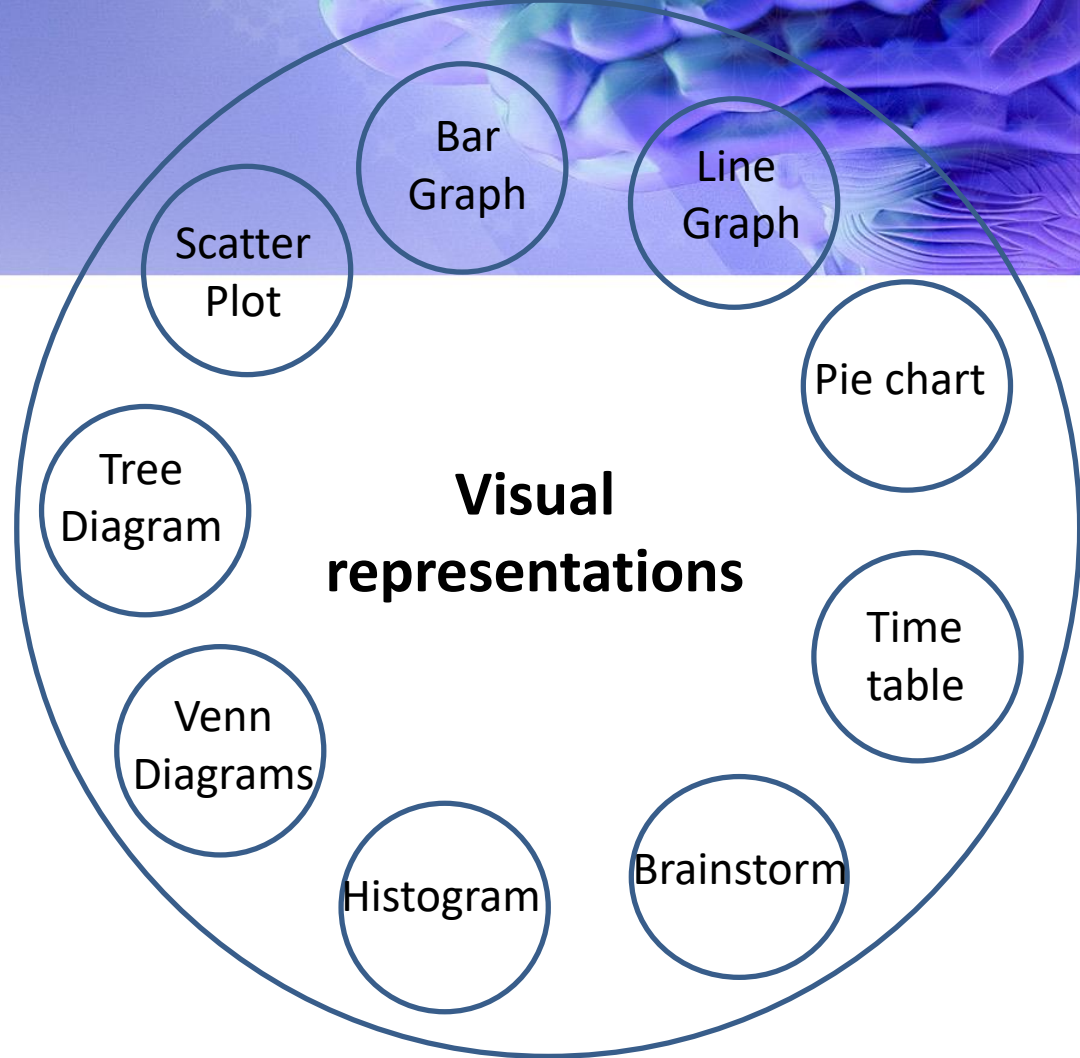
## Why to Explore?

Data exploration is done by visualizing the data or arranging the data uniformly for a better understanding.

- Thus, to analyze the data, we need to visualize it in some user-friendly format so that we can:
  - Quickly get a sense of the trends, relationships and patterns contained within the data.
  - Define strategy for which model to use at a later stage.
  - Communicate the same to others effectively.

# Data Exploration:

## Data Visualization Techniques:



# Data Exploration:

Activity: List down 5 new data visualization techniques which you learnt from

<b>Data Visualisation Technique 1</b>	
Name of the Representation	
One-line Description	
How to draw it	
Suitable for which data type?	